SEMANTICS

Matt Post
IntroHLT class
15 September 2020

JOHNS HOPKINS UNIVERSITY

From last time

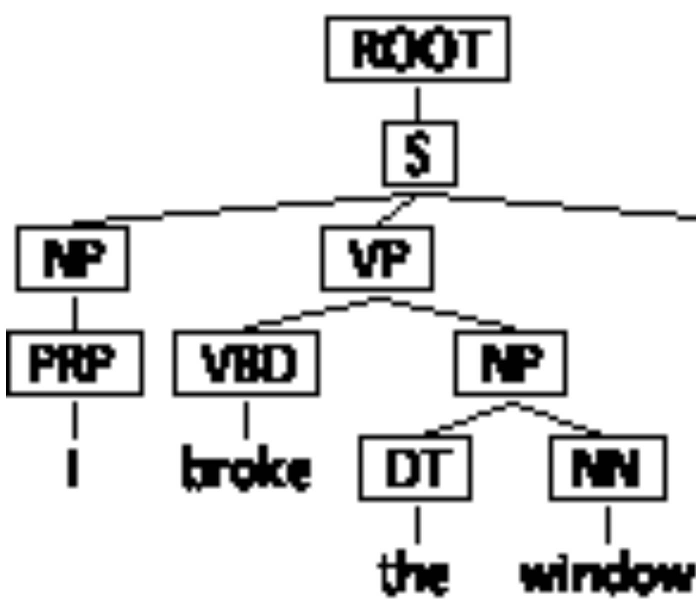
- How can we determine the core dependencies in the short sentences to the right?

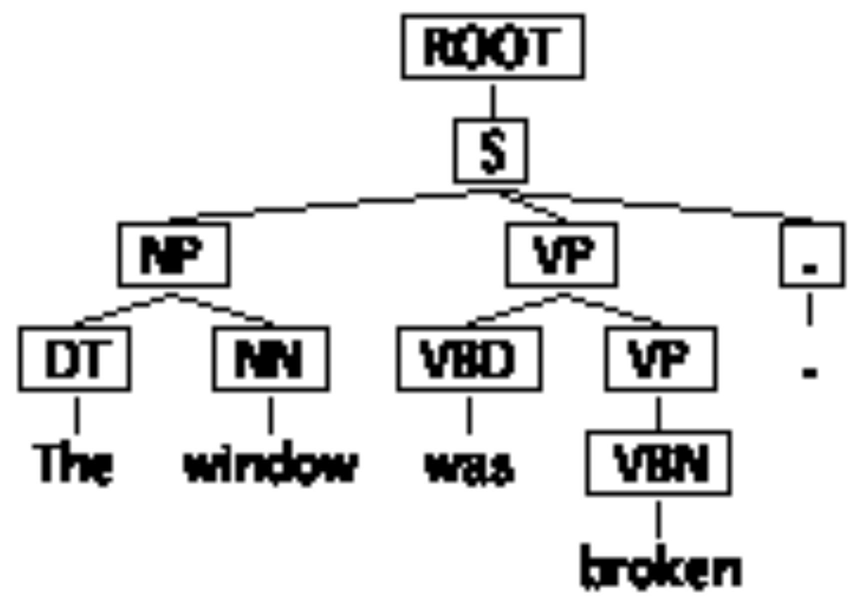
Him the Almighty hurled

Dipanjan taught Johnmark

Semantic Roles

- Syntax describes the grammatical relationships between words and phrases
  - But there are many different ways to express a particular meaning





- These variations miss an important generalization

- Structure is important, but one way it is important is as a “scaffolding for meaning”
- What we want to know is

    who *did* ***what*** *to* ***whom***

    (and ***when****)*

    (and ***where****)*

    (and ***how****)*

A linguistic hierarchy

pragmatics
semantics
syntax
morphology
phonetics

how can we represent knowledge?

how do we do so in pursuit of solving some task?

Goal

- Given a sentence
  - answer the question “**who** did **what** to **whom** etc”
  - store answer in a machine-usable way

Goal

- Given a sentence
  - answer the question “**who** did **what** to **whom** etc”
  - store answer in a machine-usable way
- This requires
  - specifying some representation for meaning
  - specifying a representation for word relationships
  - mapping the words to these representations

Goal

- Given a sentence
  - answer the question “**who** did **what** to **whom** etc”
  - store answer in a machine-usable way
- This requires
  - specifying some representation for meaning
  - specifying a representation for word relationships
  - mapping the words to these representations
- **How do we represent meaning?**

Semantics

UNTIL RECENTLY → **NOW**

- Explicit representations
- Backed by human-constructed databases and ontologies
- Feature-based models

- End-to-end
- Backed by very large collections of unstructured human text
- Neural models

lexical semantics

"Word Senses and WordNet"
https://web.stanford.edu/~jurafsky/slp3/19.pdf

Words have many meanings

- Example
  - She pays 3% **interest** on the loan.
  - He showed a lot of **interest** in the painting.
  - Microsoft purchased a controlling **interest** in Google.
  - It is in the national **interest** to invade the Bahamas.
  - I only have your best **interest** in mind.
  - Playing chess is one of my **interests**.
  - Business **interests** lobbied for the legislation.

Words overlap in meaning

- What is the relationship among these words?
  - *{organization, team, group, association, conglomeration, institution, establishment, consortium, federation, agency, coalition, alliance, league, club, confederacy, syndicate, society, corporation}*
  - organisation?

Word senses can be organized

- **Synset**: a group of words with a shared meaning
  - This generalizes the notion of a word
  - Nowadays we'd think of this as a cluster in some high-dimensional space
- We can then define relationships between these sets of words

Relationships

- Many-many relationship between form and meaning

Relationships

- Many-many relationship between form and meaning
- Same forms

Relationships

- Many-many relationship between form and meaning
- Same forms
  - **polysemy**   many related meanings

Relationships

- Many-many relationship between form and meaning
- Same forms
  - **polysemy** many related meanings
  - **homonymy** different meanings

Relationships

- Many-many relationship between form and meaning
- Same forms
  - **polysemy** many related meanings
  - **homonymy** different meanings
- Different forms

Relationships

- Many-many relationship between form and meaning
- Same forms
  - **polysemy** many related meanings
  - **homonymy** different meanings
- Different forms
  - **synonymy** same / similar meanings

Relationships

- Many-many relationship between form and meaning
- Same forms
  - **polysemy** many related meanings
  - **homonymy** different meanings
- Different forms
  - **synonymy** same / similar meanings
  - **antonymy** opposite or contrary meaning

More relationships

- Hypernym / hyponym
  - IS-A(animal, cat)
  - cat → feline → carnivore → placental mammal → mammal → vertebrate → …
- Meronymy (part / whole)
  - HAS-PART(cat, paw)
  - IS-PART-OF(paw, cat)
- Membership
  - IS-MEMBER-OF(professor, faculty)
  - HAS-MEMBER(faculty, professor)

WordNet

- English WordNet: https://wordnet.princeton.edu/

WordNet

- English WordNet: https://wordnet.princeton.edu/
  - nouns, verbs adjectives

WordNet

- English WordNet: https://wordnet.princeton.edu/
  - nouns, verbs adjectives
- Multilingual WordNet: http://compling.hss.ntu.edu.sg/omw/

WordNet

- English WordNet: https://wordnet.princeton.edu/
  - nouns, verbs adjectives
- Multilingual WordNet: http://compling.hss.ntu.edu.sg/omw/
- Examples: interest, tiger

Example (interest)

WordNet link

Noun

- S: (n) **interest**, involvement (a sense of concern with and curiosity about someone or something) *"an interest in music"*
- S: (n) sake, **interest** (a reason for wanting something done) *"for your sake"; "died for the sake of his country"; "in the interest of safety"; "in the common interest"*
- S: (n) **interest**, interestingness (the power of attracting or holding one's attention (because it is unusual or exciting etc.)) *"they said nothing of great interest"; "primary colors can add interest to a room"*
- S: (n) **interest** (a fixed charge for borrowing money; usually a percentage of the amount borrowed) *"how much interest do you pay on your mortgage?"*
- S: (n) **interest**, stake ((law) a right or legal share of something; a financial involvement with something) *"they have interests all over the world"; "a stake in the company's future"*
- S: (n) **interest**, interest group ((usually plural) a social group whose members control some field of activity and who have common aims) *"the iron interests stepped up production"*
- S: (n) pastime, **interest**, pursuit (a diversion that occupies one's time and thoughts (usually pleasantly)) *"sailing is her favorite pastime"; "his main pastime is gambling"; "he counts reading among his interests"; "they criticized the boy for his limited pursuits"*

Example (synset)

(a person who is gullible and easy to take advantage of)

S: (n) chump, fool, gull, mark, patsy, fall guy, sucker, soft touch, mug (a person who is gullible and easy to take advantage of)

Jurafsky & Martin, 3rd Ed., Ch 19. p. 6

Word Sense Disambiguation

- How can we map word (tokens) to the correct sense?

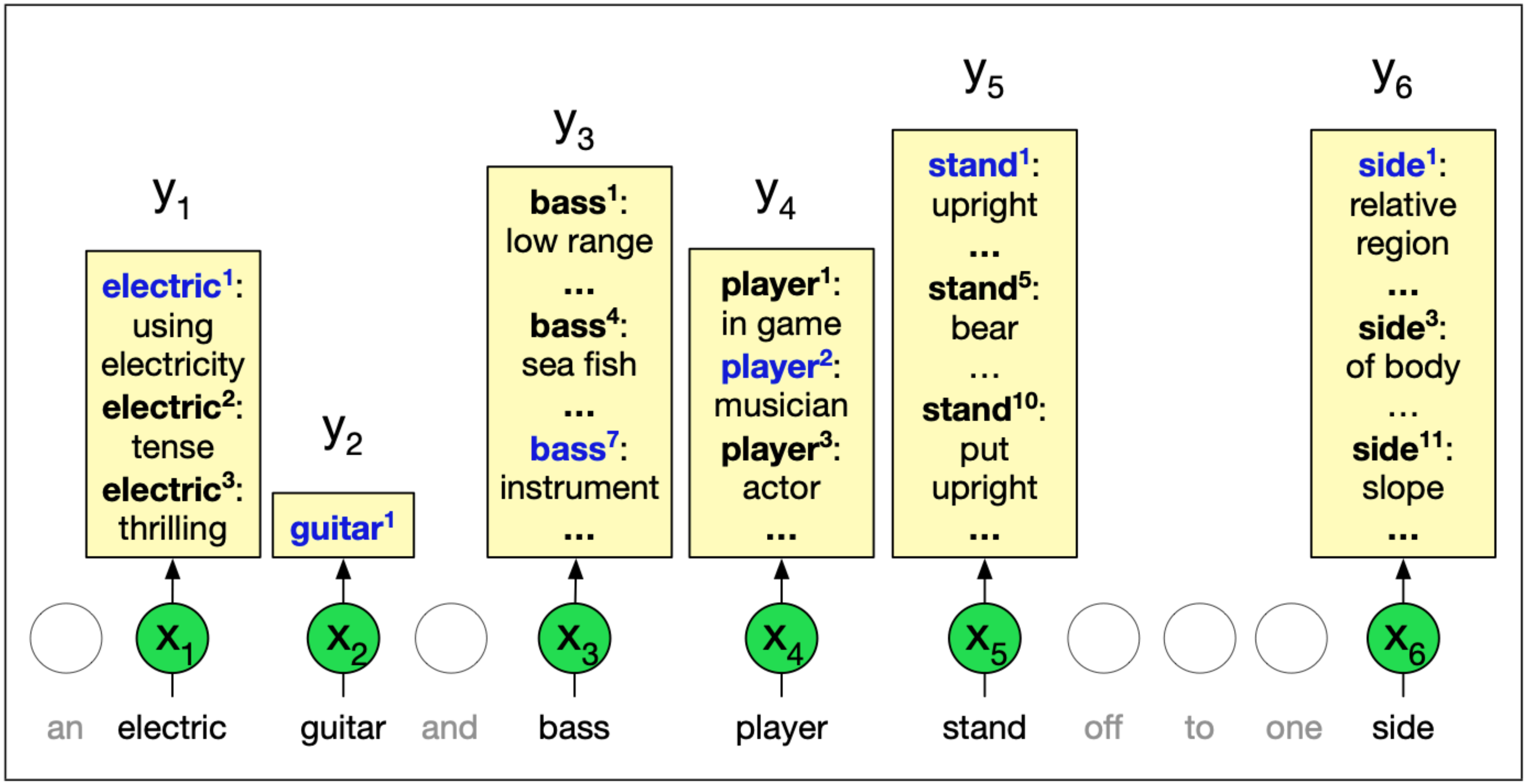


Figure 19.8 The all-words WSD task, mapping from input words (x) to WordNet senses (y). Only nouns, verbs, adjectives, and adverbs are mapped, and note that some words (like *guitar* in the example) only have one sense in WordNet. Figure inspired by Chaplot and Salakhutdinov (2018).

Supervised WSD

- Supervised approach
  - Take a corpus tagged with senses
  - Train a model on these tags
  - Apply to new data at test time
  -

Feature-based models

- Define features that are predictive of senses
  - window of words around the word
  - POS tags of window words
  - parse tree features
  - …you get the picture
- Learn a model using standard ML techniques, typically
  - P(sense | word, features)
  - e.g., maxent, naive Bayes, CRF

Contextual Embeddings

- The modern approach
- Compute contextual embeddings using (say) BERT or ELMo over a labeled dataset
  - produce a cluster by averaging the embeddings over the whole (labeled) training data
  - this produces a cluster for every sense of a word
  - at test time, again compute the contextual embedding, then assign by nearest-neighbors

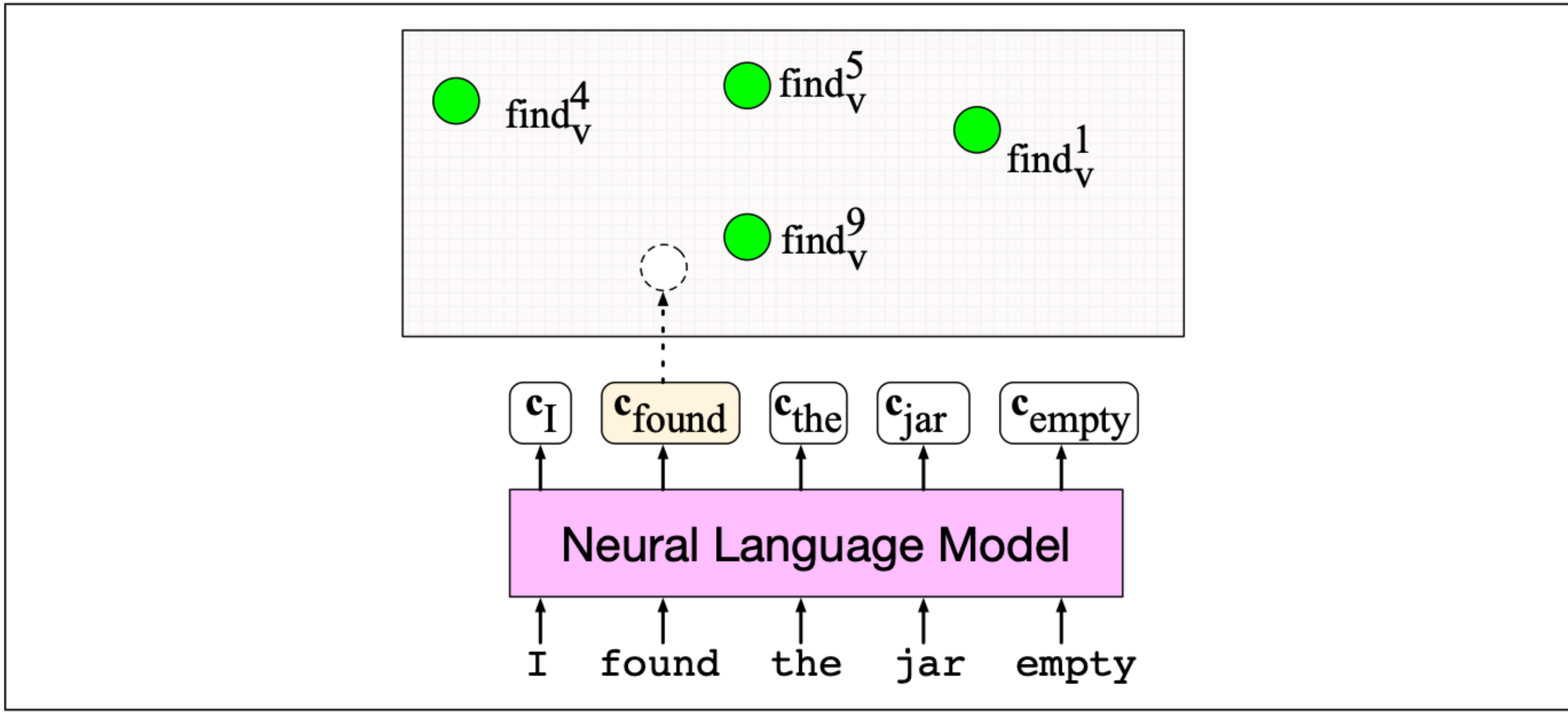


Figure 19.9 The nearest-neighbor algorithm for WSD. In green are the contextual embeddings precomputed for each sense of each word; here we just show a few of the senses for *find*. A contextual embedding is computed for the target word *found*, and the and then the nearest neighbor sense (in this case $\mathbf{find}_n^9$) would be chosen. Figure inspired by Loureiro and Jorge (2019).

Unsupervised WSD

- Consider:
  - we have Wordnet
    - which has groups of word forms, along with a gloss or definition
      - organized hierarchically
- What if you don’t have labeled data to choose from? How might you assign the correct word sense?

Lesk Algorithm

- (19.19) “The bank can guarantee deposits will eventually cover future tuition costs because it invests in adjustable-rate mortgage securities.”

bank^1	Gloss:	a financial institution that accepts deposits and channels the money into lending activities
	Examples:	“he cashed a check at the bank”, “that bank holds the mortgage on my home”
bank^2	Gloss:	sloping land (especially the slope beside a body of water)
	Examples:	“they pulled the canoe up on the bank”, “he sat on the bank of the river and watched the currents”

  - which is the correct assignment?

Other approaches

- WordNet is huge, complicated, expensive to build
- Clustering
  - Instead of mapping words to predefined senses, using clustering algorithms to induce unlabeled clusters
  - Compute cluster centroids
  - At test time, assign words to clusters based on the nearest centroid
  - This has obvious connections to **word embeddings**

Summary

- Some takeaways:
  - Words can be grouped according to their overlapping senses called **synsets**
  - These groups can then be organized into an ontology with relationships
  - WordNet is a large database of these synsets, primarily for English
- Further reading:
  - Jurafsky & Martin, 3rd Ed., Chapter 19
    https://web.stanford.edu/~jurafsky/slp3/19.pdf

semantic role labeling

Semantic Role Labeling

- Assuming we can disambiguate a word, can we get back to the core question of identifying word relationships?
- Example sentence pair from before
  - *I broke the window*
  - *The window was broken by me*
- There is a generalization here involving the types of participants

Much of the structure here follows Chapter 20 of Jurafsky & Martin, 3rd Ed.

https://web.stanford.edu/~jurafsky/slp3/20.pdf

Thematic Roles

Thematic Role	Definition
AGENT	The volitional causer of an event
EXPERIENCER	The experiencer of an event
FORCE	The non-volitional causer of the event
THEME	The participant most directly affected by an event
RESULT	The end product of an event
CONTENT	The proposition or content of a propositional event
INSTRUMENT	An instrument used in an event
BENEFICIARY	The beneficiary of an event
SOURCE	The origin of the object of a transfer event
GOAL	The destination of an object of a transfer event

Figure 20.1 Some commonly used thematic roles with their definitions.

Thematic Roles

Thematic Role	**Example**
AGENT	*The waiter* spilled the soup.
EXPERIENCER	*John* has a headache.
FORCE	*The wind* blows debris from the mall into our yards.
THEME	Only after Benjamin Franklin broke *the ice*...
RESULT	The city built a *regulation-size baseball diamond*...
CONTENT	Mona asked *"You met Mary Ann at a supermarket?"*
INSTRUMENT	He poached catfish, stunning them *with a shocking device*...
BENEFICIARY	Whenever Ann Callahan makes hotel reservations *for her boss*...
SOURCE	I flew in *from Boston*.
GOAL	I drove *to Portland*.

Figure 20.2 Some prototypical examples of various thematic roles.

FrameNet

- **frame**: the general background information relating to an event that is invoked and filled by the sentence
  - established idea in cognitive science and semantics
  - related to the idea of **scripts** (story patterns that underly an event or report)

Example

- Consider these sentences

(20.20) [$_{\text{ITEM}}$ Oil] *rose* [$_{\text{ATTRIBUTE}}$ in price] [$_{\text{DIFFERENCE}}$ by 2%].

(20.21) [$_{\text{ITEM}}$ It] has *increased* [$_{\text{FINAL\_STATE}}$ to having them 1 day a month].

(20.22) [$_{\text{ITEM}}$ Microsoft shares] *fell* [$_{\text{FINAL\_VALUE}}$ to 7 5/8].

(20.23) [$_{\text{ITEM}}$ Colon cancer incidence] *fell* [$_{\text{DIFFERENCE}}$ by 50%] [$_{\text{GROUP}}$ among men].

(20.24) a steady *increase* [$_{\text{INITIAL\_VALUE}}$ from 9.5] [$_{\text{FINAL\_VALUE}}$ to 14.3] [$_{\text{ITEM}}$ in dividends]

(20.25) a [$_{\text{DIFFERENCE}}$ 5%] [$_{\text{ITEM}}$ dividend] *increase*...

these can be thought of as invoking the following frame

Core Roles	
ATTRIBUTE	The ATTRIBUTE is a scalar property that the ITEM possesses.
DIFFERENCE	The distance by which an ITEM changes its position on the scale.
FINAL_STATE	A description that presents the ITEM's state after the change in the ATTRIBUTE's value as an independent predication.
FINAL_VALUE	The position on the scale where the ITEM ends up.
INITIAL_STATE	A description that presents the ITEM's state before the change in the ATTRIBUTE's value as an independent predication.
INITIAL_VALUE	The initial position on the scale from which the ITEM moves away.
ITEM	The entity that has a position on the scale.
VALUE_RANGE	A portion of the scale, typically identified by its end points, along which the values of the ATTRIBUTE fluctuate.
Some Non-Core Roles	
DURATION	The length of time over which the change takes place.
SPEED	The rate of change of the VALUE.
GROUP	The GROUP in which an ITEM changes the value of an ATTRIBUTE in a specified way.

Figure 20.3 The frame elements in the **change_position_on_a_scale** frame from the FrameNet Labelers Guide (Ruppenhofer et al., 2016).

Semantic Role Labeling: the task

- Determine semantic roles of words in a sentence
  - Input: *You can't blame the program for being unable to identify it.*
  - Output: [You]$_{\textbf{COGNIZER}}$ can't [blame]$_{\textbf{TARGET}}$ [the program]$_{\textbf{EVALUEE}}$ [for being unable to identify it]$_{\textbf{REASON}}$

The algorithm

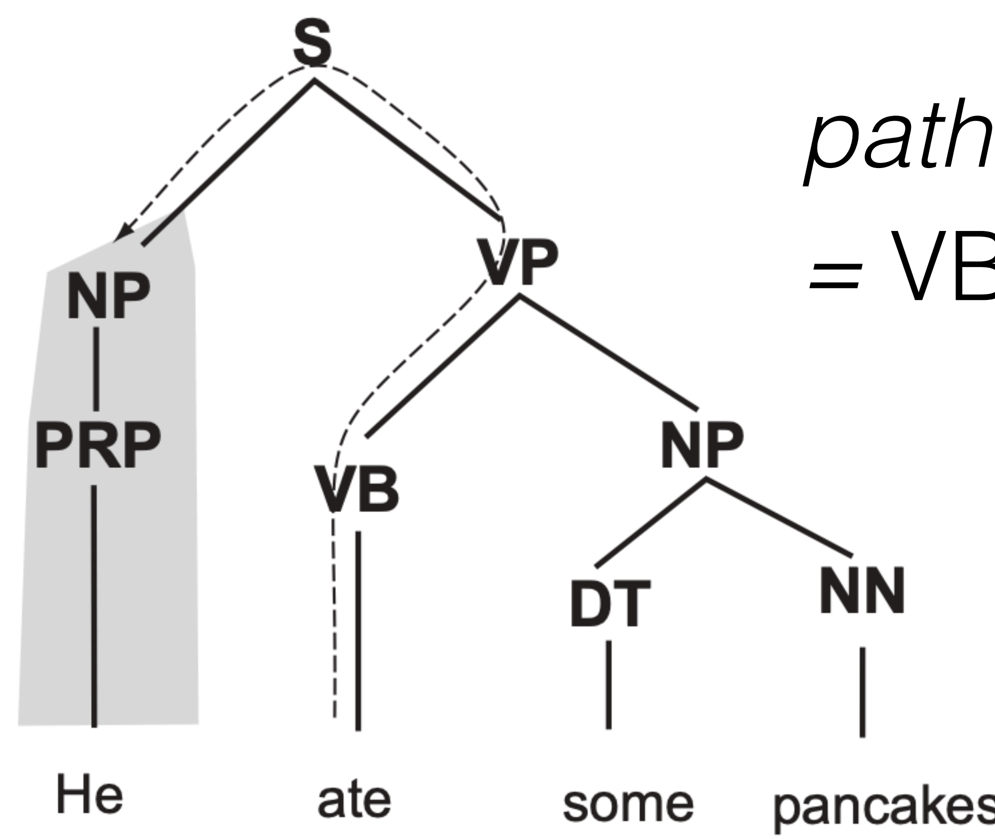
```
function SEMANTICROLELABEL(words) returns labeled tree

    parse ← PARSE(words)
    for each predicate in parse do
        for each node in parse do
            featurevector ← EXTRACTFEATURES(node, predicate, parse)
            CLASSIFYNODE(node, featurevector, parse)
```

Figure 20.4 A generic semantic-role-labeling algorithm. CLASSIFYNODE is a 1-of-*N* classifier that assigns a semantic role (or NONE for non-role constituents), trained on labeled data such as FrameNet or PropBank.

Features

- Nonterminal label ("NP")
- Governing category ("S" or "VP" = subject or object)
- Parse tree path
- Position (before or after predicate)
- Head word
- Many, many more



path(ate→He)

$= VB \uparrow VP \uparrow S \downarrow NP$

- Trained with discriminative ML algorithms (SVM, MaxEnt)

Bringing it together

- This can finally bring us to the point where we have tuples, say of the form (action, agent, patient, [theme])
  - e.g., (saw, man, bird, telescope)
- How can we use these?

Bringing it together

- This can finally bring us to the point where we have tuples, say of the form (action, agent, patient, [theme])
  - e.g., (saw, man, bird, telescope)
- How can we use these?
- Maybe question answering:
  - build large database of tuples
  - for a new question:
    - map it to a tuple
    - match it against the database, fill in the slot

Use case: MT evaluation

- Task: determine the quality of an **MT system output** by comparing it against a **human reference**
- *MEANT: An inexpensive, high-accuracy, semi-automatic metric for evaluating translation utility based on semantic roles* (**Lo & Wu**, ACL 2011)

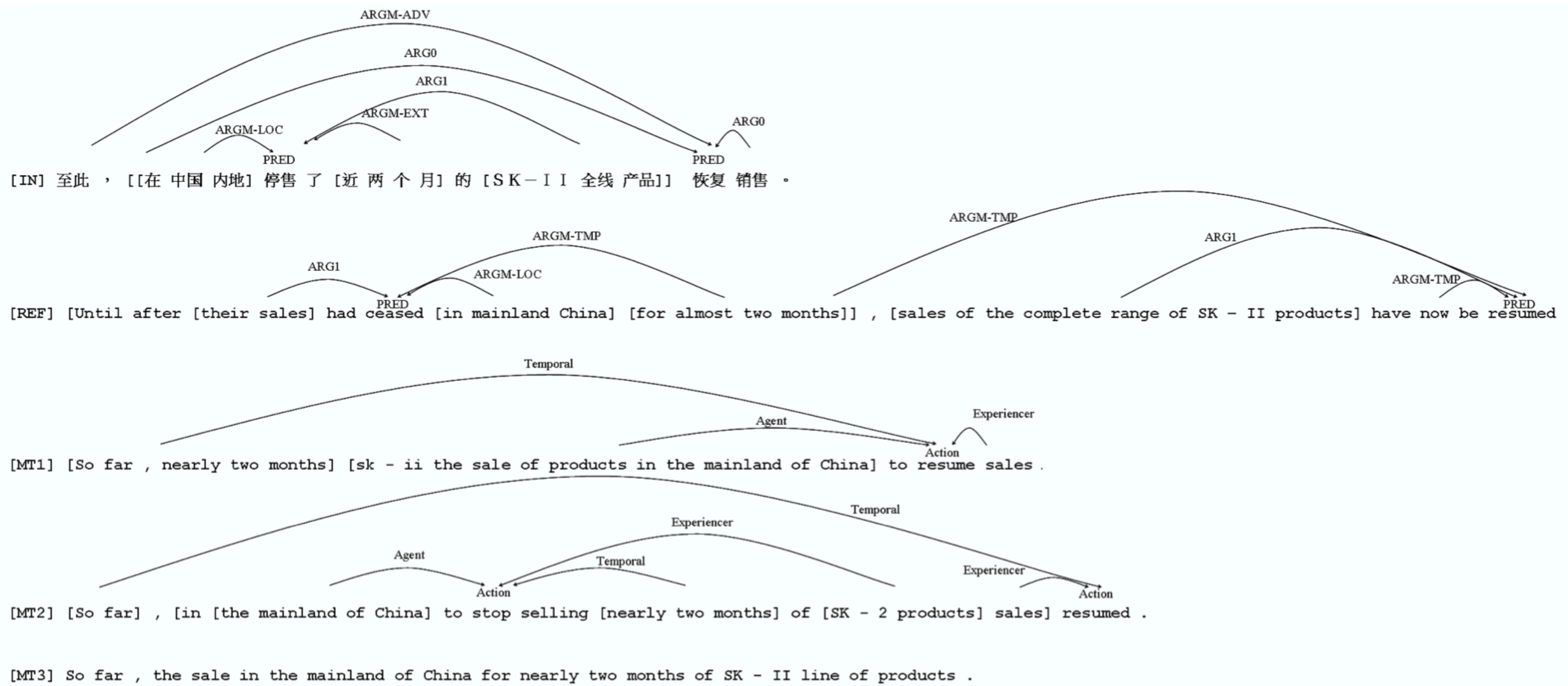


Figure 1: Example of source sentence and reference translation with reconstructed semantic frames in Propbank format and MT output with reconstructed semantic frames by minimal trained human annotators. Following Propbank, there are no semantic frames for MT3 because there is no predicate.

MEANT results

- When used to compare two system outputs, MEANT performs well (compared to other automatic metrics) in ranking them the same way humans would

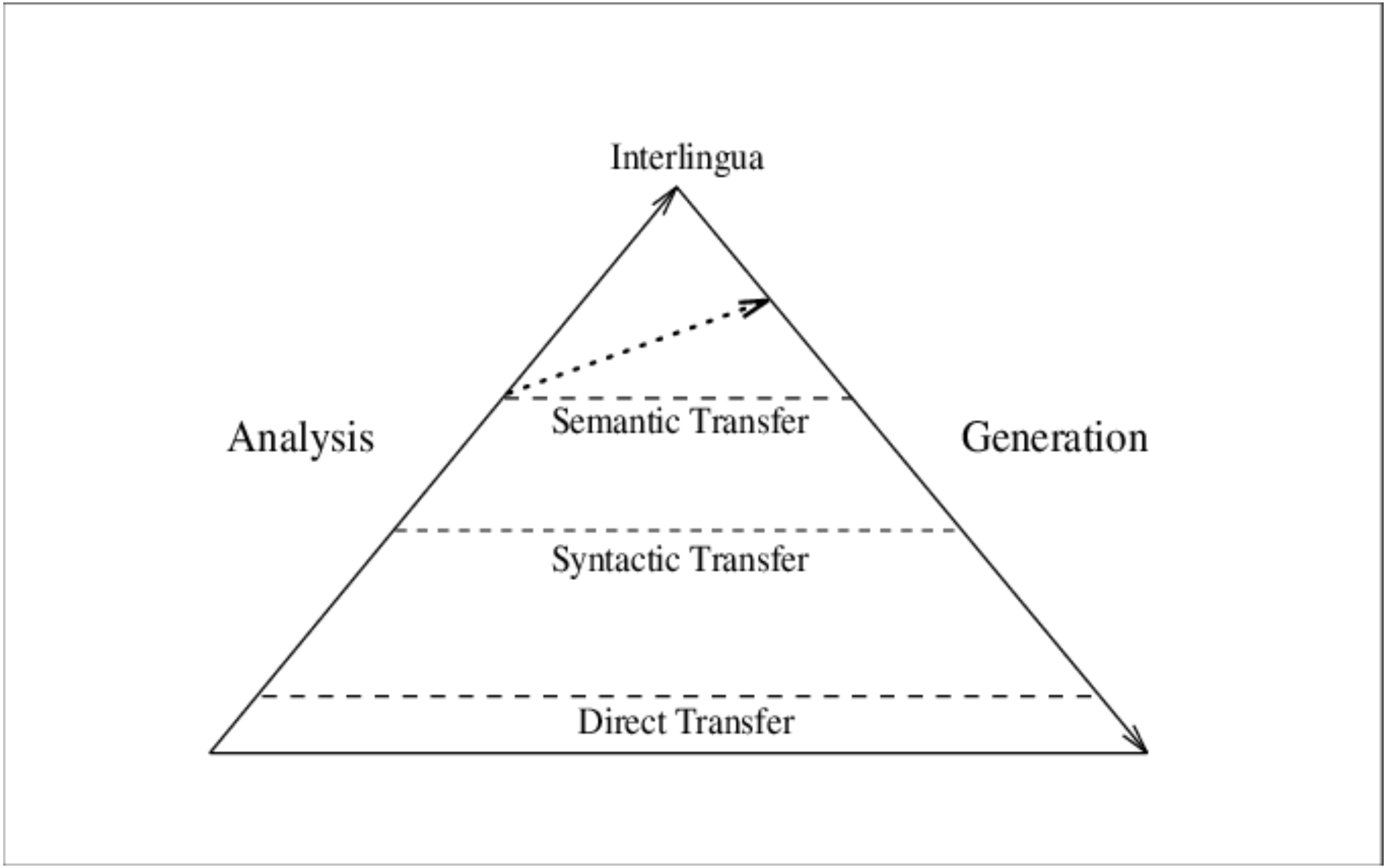
Table 8: Sentence-level correlation with human adequacy judgments. *The weights for individual roles in the metric are tuned by optimizing the correlation.

Metrics	**Kendall τ**
HTER	0.4324
HMEANT gold - monolinguals *	0.4324
HMEANT auto - monolinguals *	0.3964
MEANT gold - auto *	0.3694
MEANT auto - auto *	0.3423
NIST	0.2883
BLEU / METEOR / TER / PER	0.1982
CDER	0.1171
WER	0.0991

Reflection

- Many of the representations we've considered are aimed at constructing an **explicit** representation of word and sentence meanings
- Does this make sense?

The Vauquois Triangle



Should we expect explicit representations?

Summary

- We have seen an overview of symbolic reps. of meaning
  - Extensive statistical and machine learning techniques were used to solve them
  - These dominated research for the past few decades until ~2015
- It is unclear how to represent meaning in a manner that can be operationalized
- Today's end-to-end deep learning approaches perform well without them
  - However, work continues into building systems for various tasks that combine "dumb text" with human-curated resources